Carleman Lifting for Nonlinear System Identification with Guaranteed Error Bounds

Moad Abudia[1], Joel A. Rosenfeld[2], and Rushikesh Kamalapurkar[1]

***Abstract*— This paper concerns identification of uncontrolled or closed loop nonlinear systems using a set of trajectories that are generated by the system in a domain of attraction. The objective is to ensure that the trajectories of the identified systems are close to the trajectories of the real system, as quantified by an error bound that is *prescribed a priori*. A majority of existing methods for nonlinear system identification rely on techniques such as neural networks, autoregressive moving averages, and spectral decomposition that do not provide systematic approaches to meet pre-defined error bounds. The developed method is based on Carleman linearization-based lifting of the nonlinear system to an infinite dimensional linear system. The linear system is then truncated to a suitable order, computed based on the prescribed error bound, and parameters of the truncated linear system are estimated from data. The effectiveness of the technique is demonstrated by identifying an approximation of the Van der Pol oscillator from data within a prescribed error bound.**

I. Introduction

Identifying nonlinear dynamical systems from data, without knowledge of the structure of the model, and with guaranteed error bounds, has proven to be a difficult challenge. The focus of this paper is on the identification of the dynamic model of a nonlinear system within prescribed error bounds using Carleman lifting, where the finite dimensional nonlinear system is lifted into an infinite-dimensional linear system via Carleman linearization. The linear system is then truncated to a suitable order, computed based on the prescribed error bound, and the parameters of the truncated linear system are estimated from data.

A variety of system identification methods are available for linear time invariant (LTI) systems. For example, in non-parametric frequency-domain estimation methods, the system is subjected to a white noise input, the output is represented in the frequency domain, and the user fits a model that is determined by the number of resonant frequencies and the decay rate of the power spectrum. On the other hand, parametric methods presuppose a system representation, such as an auto regressive moving average (ARMA) model or a state space (SS) model, where the order of the system is estimated first using auto correlation. The system identification problem is reduced to a parameter estimation problem, which is solved using numerical gradient descent methods to minimize a suitable error metric (c.f. [1]).

In the case of nonlinear systems, the identification problem is much more difficult. If the structure of the system is known *a priori*, the identification problem can be reduced to a parameter estimation problem which can be solved using Lyapunov-based adaptive estimation methods [2]. For problems with unknown system structure, spectral decomposition methods have proven to be effective for system identification in many practical applications [3]–[5], they generally do not provide theoretical error bound guarantees. Similarly, using neural networks for nonlinear system identification [6], do not provide theoretical error bound guarantees.

In [7], the authors develop a system identification approach that bounds the model error in terms of the prediction horizon and the magnitude of the derivatives of the system state, if the magnitudes of the derivatives of the system state are not analytically available, then they are calculated from data, in which case the error bounds may be violated.

The approach in this paper is inspired by the Carleman linearization approach developed in [8]. Carleman linearization converts a finite dimensional nonlinear system to an infinite dimensional linear system. The conversion is realized by expanding the state space of the nonlinear system to include all monomial functions of the original state variables.

In [9], the authors used Carleman linearization in conjunction with a known nonlinear model to develop a truncated linear system that approximates the nonlinear system in a region of attraction. Moreover, the trajectory of the truncated linear system is guaranteed to stay within a computable error bound around the trajectory of the nonlinear system over a prediction horizon. The error bound can be computed *a priori* when the decay rate of the Maclaurin expansion of the nonlinear system and the order of truncation are known.

In this paper, the results in [9] are leveraged to develop a Carleman lifting-based approach to data-driven modeling of nonlinear systems. Given a set of trajectories generated by a nonlinear system and an upper bound on the decay rate of the Maclaurin expansion, a linear system of a higher order is identified such that the trajectory of the identified linear system stays within a guaranteed error bound around the trajectory of the nonlinear system over a prediction horizon, and the error bound can be *prescribed a priori*.

The paper is organized as follows. The problem is for-

[1] The authors are with the School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK, USA. {abudia, rushikesh.kamalapurkar}@okstate.edu.

[2] The author is with the Department of Mathematics and Statistics, University of South Florida, Tampa, Fl, USA. rosenfeldj@usf.edu.

This research was supported, in part, by the National Science Foundation (NSF) under award number 2027999, and the Air Force Office of Scientific Research under award number FA9550-20-1-0127. Any opinions, findings, conclusions, or recommendations detailed in this article are those of the author(s), and do not necessarily reflect the views of the sponsoring agencies.

mulated in Section II. In Section III the approach for identification of linear systems is presented. In Section IV the properties of the identified system are analyzed. In Section V an algorithm to generate the identified linear system with a guaranteed error bound is presented. In Section VI a simulation is presented to demonstrate the system identification method. In Section VII, a discussion of the results is presented, and Section VIII includes concluding remarks and a discussion on future work.

II. Problem Statement

Consider an unknown dynamical system of the form

$$\dot{x} = \mathbf{f}(t, x) \tag{1}$$

where $\mathbf{f} : \left(\mathbb{R}_+ \times \mathbb{R}^d\right) \to \mathbb{R}^d$ is a vector valued real analytic function of several variables, $\mathbb{R}_+ \coloneqq \{y \in \mathbb{R} : y \geq 0\}$, $x(0) = x_0$, and $f(t, 0) = 0$. Given a set of observed trajectories, $\{\gamma_i\}_{i=1}^m = X$, generated from (1) and an error bound $\Delta > 0$, the objective is to develop a systematic technique to either

1) construct another dynamical system of the form $\dot{\hat{z}} = g(t, \hat{z})$ such that the error $\|x(t) - \hat{z}|_d(t)\|$ between the trajectories of the $\hat{z}-$system and the trajectories of (1) is less than Δ for $t \in [0, \tau^*]$, for some $\tau^* > 0$, where $\hat{z}|_d$ is the truncation of $\hat{z}$ to the first d dimensions, or
2) conclude that construction of such a system, using the particular method developed in this paper, is not possible.

The problem, as formulated above, is difficult to solve for general nonlinear systems. In this paper the formulation is restricted to a sub-class of nonlinear systems, defined by the following assumptions as in [9].

Assumption 1. The vector field $\mathbf{f}(t, x)$ admits a Maclaurin expansion about the state vector x

$$\mathbf{f}(t, x) = \sum_{\boldsymbol{\alpha} \in \mathbb{Z}_+^d} \mathbf{f}_{\boldsymbol{\alpha}}(t)\mathbf{x}^{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha} \in \mathbb{Z}_+^d \setminus \{0\}} \mathbf{f}_{\boldsymbol{\alpha}}(t)\mathbf{x}^{\boldsymbol{\alpha}}, t \in \mathbb{R}^+. \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_+^d$, is a multi-index and its corresponding multi-variate monomial is $\mathbf{x}^{\alpha} = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, where $\mathbb{Z}_+^d$ is the set of non-negative integers.

Assumption 2. The Maclaurin expansion coefficients satisfy the exponential decay property

$$\sup_{t \geq 0} \sum_{|\alpha| = n} \|\mathbf{f}_{\boldsymbol{\alpha}}(t)\|_{\infty} \leq CR^{-n}, n \geq 0 \tag{3}$$

where C and R are positive constants, and the cardinality of $\boldsymbol{\alpha}$ is defined by $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d$.

While this assumption is restrictive, a large subclass of systems fall under this category, since polynomial, trigonometric, and exponential functions have Maclaurin expansions with exponential decay. The idea is to use Carleman lifting [8], [9] to lift the nonlinear system into an infinite-dimensional linear system such that the truncation of the trajectory of the linear system to the first d dimensions approximates the trajectory of the system in (1), in the infinity norm over a finite time interval. The infinite-dimensional linear system is then truncated to yield a finite-dimensional linear system such that the error between the projected trajectories and the original trajectories is less than the given error bound, Δ.

III. Data-driven Carleman Lifting

Carleman linearization [8] lifts a finite dimensional nonlinear system to an infinite dimensional linear system

$$\dot{y} = \mathcal{A}y \tag{4}$$

where y is an infinite dimensional vector consisting of all unrepeated monomials of x and $\mathcal{A}$ is an infinite dimensional operator. The operator $\mathcal{A}$ is approximated by truncating its matrix representation to the first $\mathbb{M} \times \mathbb{M}$ block, which will be called A.

The nonlinear system is then approximated by a linear system in terms of a lifted state z which consists of unrepeated monomials of x up to order N. For example, if $N = 3$ then

$$z = l(x) = \begin{array}{c}[x_1, \ldots, x_d, x_1^2, x_1x_2, \ldots, x_1x_d, x_2^2, \ldots, x_d^2, \\ x_1^3, x_1^2x_2, \ldots, x_1x_2x_3, \ldots, x_d^3]^T,\end{array}$$

where $l : \mathbb{R}^d \to \mathbb{R}^{\mathbb{M}}$ denotes the lifting map and $\mathbb{M}$ is the number of monomials which is $\mathbb{M} = \sum_{k=1}^{N} \binom{k+d-1}{d-1}$.

The results of [9] indicate that for any system that satisfies Assumptions 1 and 2, a trajectory in a region of attraction can be approximated, with arbitrary accuracy, by truncation of solutions of

$$\dot{z} = Az \tag{5}$$

to the first d dimensions, where computation of A requires complete knowledge of the system dynamics, $\mathbf{f}$.

In this paper, a data-driven approach to generate the linear system is developed. Given a set of trajectories of the nonlinear system, denoted by $\{\gamma_i\} : [0, T] \to \mathbb{R}^d$, that satisfy $\|\gamma_i(t)\| \leq M$ for all $t \in [0, T]$, the objective is to find an order N and an estimate $\hat{A}$ of an N^{th}-order truncation A of the operator $\mathcal{A}$ such that $\sup_{t \in [0, \tau^*]} \|x(t) - \hat{z}|_d(t)\| \leq \Delta$, where $[0, \tau^*]$ is the interval over which the error bound can be guaranteed and $\hat{z}|_d$ is the $d-$dimensional truncation of the solution $\hat{z}$ of $\dot{\hat{z}} = \hat{A}\hat{z}$ starting from $\hat{z}(0) = l(x(0))$.

Let z_i denote the trajectory of (5) starting from the initial condition $z_i(0) = l(\gamma_i(0))$. Let

$$\epsilon_i(t) \coloneqq l(\gamma_i(t)) - z_i(t) \tag{6}$$

denote the error between the trajectories of the model-based Carleman linearization in (5) and the trajectories of (1). The trajectories of the linear and the nonlinear systems are then related by

$$\frac{d}{dt} l(\gamma_i(t)) = \dot{z}_i(t) + \dot{\epsilon}_i(t) = Az_i(t) + \dot{\epsilon}_i(t). \tag{7}$$

Integrating (7),

$$\int_0^T \frac{d}{dt} l(\gamma_i(t)) = \int_0^T A z_i(\tau) d\tau + \epsilon_i(T)$$
$$= A \int_0^T \left(l(\gamma_i(\tau)) - \epsilon_i(\tau) \right) d\tau + \epsilon_i(T). \quad (8)$$

Note that $\epsilon_i(0) = 0$. Concatenating all the measured trajectories into a vector $\Gamma(t) = [l(\gamma_1(t)), \ldots, l(\gamma_m(t))]^\top$, and letting $I = \int_0^T Z(\tau) d\tau$, where $Z = [z_1, \ldots, z_m]$, with $I_\epsilon = \int_0^T \epsilon(\tau) d\tau$ and $I_\Gamma = \int_0^T \Gamma(\tau) d\tau$, one gets the relationships $I_\Gamma = I + I_\epsilon$ and

$$\Gamma(T) - \Gamma(0) = A I_\Gamma - A I_\epsilon + \epsilon(T), \quad (9)$$

where $\epsilon(t) = [\epsilon_1(t), \ldots, \epsilon_m(t)]^\top$.

Provided the matrix I_Γ is full rank, a data-driven approximation $\hat{A}$ of the lifted matrix A can be computed using the least-squares solution of (9) as

$$\hat{A} = (\Gamma(T) - \Gamma(0)) I_\Gamma^\dagger, \quad (10)$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse.

Once the matrix $\hat{A}$ is computed from data, trajectories of the linear system $\dot{\hat{z}} = \hat{A}\hat{z}$ can be computed and truncated to the first d dimensions to estimate the trajectories of the nonlinear system.

IV. Error Analysis and Prescribed Error Approximation

There are two sources of error between the trajectories of the identified system truncated to the first d dimensions, and the trajectories of the nonlinear system. The first, is the error ϵ introduced in (6), between the trajectories of the nonlinear system and the model-based Carleman-linearized system in (5). The second, is the error $\tilde{z}(t) := z(t) - \hat{z}(t)$, between the trajectories of the model-based Carleman-linearized system and the identified linear system, starting from the same initial conditions. The former is quantified in [9] as follows.

Assumption 3. The measured trajectories are bounded such that $\|\gamma_i(t)\| \leq M$ for all $i = 1, ..., d$ and $t \in [0, T]$ and $M < \frac{R}{e}$, where R is introduced in (3) and e is the base of the natural logarithm.

Under Assumptions 1 - 3, for every $N \geq 1$ there exists a $\tau^* > 0$ such that

$$\|\epsilon|_d(t)\| \leq D\mu^N, \forall t \in [0, \tau^*], \quad (11)$$

where $\epsilon|_d$ is the $d-$dimensional truncation of ϵ in (6), $0 < D = M\left(1 - \left(\frac{M}{R}\right)\right)^{-1}$, $C_0 \leq CR^{-1}$, and $\mu < \left(MeR^{-1}\right) e^{C_0 \tau^*} < 1$. The error bound in (11) can then be guaranteed over the time interval $[0, \tau^*]$ for all τ^* such that

$$\tau^* < \frac{-\log(MeR^{-1})}{C_0}, \quad (12)$$

One of the main contribution of this paper is to quantify the error $\tilde{z}$, between the trajectories of the identified system and the model-based Carleman-linearized system. To that end, the following assumption is needed,

Assumption 4. There exists a constant $\overline{I}$ such that $\left\|(I^\top I)^{-1}\right\| \leq \overline{I}$, where $I = \int_0^T Z(\tau) d\tau$.

Theorem 1. *If Assumptions 1 - 4 hold, then the state estimation error is bounded as*

$$\|x(t) - \hat{z}|_d(t)\| \leq D\mu^N + \tau^* \overline{B} \overline{z} \overline{A},$$

for all $t \in [0, \tau^*]$, *where* $\overline{B} = \max\left\{\left\|e^{A\tau^*}\right\|, 1\right\}$, $\overline{z} = \sup_{t \in [0, \tau^*]} \|\hat{z}(t)\|$, *and* $\overline{A} = \left\|A - \hat{A}\right\|$.

Proof. By using the triangle inequality

$$\|x(t) - \hat{z}|_d(t)\| \leq \|\epsilon|_d(t)\| + \|\tilde{z}\|, \quad (13)$$

and under Assumptions 1 - 3.

$$\|\epsilon|_d(t)\| \leq D\mu^N \quad (14)$$

The dynamics of the error are given by

$$\dot{\tilde{z}} = Az - \hat{A}\hat{z}.$$

By adding and subtracting $A\hat{z}$,

$$\dot{\tilde{z}} = A\tilde{z} + (A - \hat{A})\hat{z}.$$

Using the variation of constants formula [10]

$$\tilde{z}(\tau^*) = e^{A\tau^*}\tilde{z}(0) + \int_0^{\tau^*} e^{A(\tau^* - \tau)}(A - \hat{A})\hat{z}(\tau) d\tau.$$

The triangle inequality and the Cauchy-Schwarz inequality then result in the bound

$$\|\tilde{z}(\tau^*)\| \leq \int_0^{\tau^*} \left\|e^{A(\tau^* - \tau)}\right\| \left\|A - \hat{A}\right\| \|\hat{z}(\tau)\| d\tau.$$

Since $\tilde{z}(0) = 0$, then

$$\|\tilde{z}(\tau^*)\| \leq \tau^* \overline{B} \overline{z} \overline{A}. \quad (15)$$

Since A is unknown, calculation of $\overline{A}$ requires further effort. To that end, let $D(t) = Z(t) - Z(0)$. From (5) if $II^\top$ is nonsingular, then

$$A = DI^\top (II^\top)^{-1},$$

and from (10) the estimation is calculated as

$$\hat{A} = (D + \epsilon)(I + I_\epsilon)^\top [(I + I_\epsilon)(I + I_\epsilon)^\top]^{-1}.$$

Note that the matrix I is comprised of integrals of the trajectories of the model-based Carleman-linearized system, and as such, it is unknown. However, the matrix I_Γ can be computed, and is perturbed from I by I_ϵ. Using continuity of eigenvalues of matrices with respect to elements of the matrix, it can be concluded that for small enough I_ϵ, the difference between $\left\|(I^\top I)^{-1}\right\|$ and $\left\|(I_\Gamma^\top I_\Gamma)^{-1}\right\|$ is $o(I_\epsilon)$.

The bound $\overline{A}$ can then be estimated using the matrix inverse identity [11], $(Q+V)^{-1} = Q^{-1} - Q^{-1}V(Q+V)^{-1}$, where Q and $(Q + V)$ are non-singular matrices. Applying the identity to the expression for $\hat{A}$ and using Assumption 4,

$$\overline{A} \leq \left\|\overline{I}\right\| \left(\left\|I^\top\right\| \|\epsilon\| + \left\|I_\epsilon^\top\right\| \|D\| + \left\|I_\epsilon^\top\right\| \|\epsilon\|\right)$$
$$+ \left[\left\|\overline{I}\right\| \left(\left\|I^\top\right\| \|I_\epsilon\| + \left\|I_\epsilon^T\right\| \|I\| + \|I_\epsilon\| \left\|I_\epsilon^\top\right\|\right)\right]$$
$$\cdot \left[\left\|I^\top\right\| \|D\| + \left\|I^\top\right\| \|\epsilon\| + \left\|I_\epsilon^\top\right\| \|D\| + \left\|I_\epsilon^\top\right\| \|\epsilon\|\right], \quad (16)$$

where $\left\|I^{\top}\right\| \leq \left\|I_{\Gamma}^{\top}\right\| + \left\|I_{\epsilon}^{\top}\right\|$.

Since $\overline{B} = \max\{\left\|e^{A\tau^*}\right\|, 1\}$, an estimate of $\left\|e^{A\tau^*}\right\|$ is obtained by realizing that

$$e^{A\tau^*} = e^{(A-\hat{A}+\hat{A})\tau^*} = e^{A-\hat{A}\tau^*}e^{\hat{A}\tau^*},$$

and as a result,

$$\overline{B} = \left\|e^{A\tau^*}\right\| \leq e^{\overline{A}\tau^*}\left\|e^{\hat{A}\tau^*}\right\|. \tag{17}$$

Using (17) and (16) in (15), and substituting (14) and (15) in (13), we get

$$\|x(t) - \hat{z}|_d(t)\| \leq D\mu^N + \tau^*\overline{B}\overline{z}\overline{A}, \quad \forall t \in [0, \tau^*]. \tag{18}$$

□

V. Order Selection Algorithm

Once the upper bound is calculated as a function of N using (18), a search over $N = 1, 2, \ldots, \overline{N}$ can be conducted (see Algorithm 1), to yield

$$N^* = \arg\min_N (D\mu^N + \tau^*\overline{B}\overline{z}\overline{A}),$$

where N^* denotes the lifted order which produces an identified system with the smallest guaranteed error bound. If $D\mu^{N^*} + \tau^*\overline{B}\overline{z}\overline{A} > \Delta$, then the system identification method cannot produce a system that guarantees $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\| < \Delta$. Otherwise, if $D\mu^{N^*} + \tau^*\overline{B}\overline{z}\overline{A} \leq \Delta$, then the system can be identified using the truncation order N^* to guarantee $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\| < \Delta$.

Algorithm 1 Carleman System identification algorithm. In the algorithm, $\overline{N}$ is an upper bound on the truncation order, selected a priori.

Require: X, R, C, $\overline{N}$ and Δ
Ensure: $C_0 \leq CR^{-1}$
Ensure: $M > \|x\|_\infty$
  $D \leftarrow M\left(1 - \left(\frac{M}{R}\right)\right)^{-1}$
Ensure: $D > 0$
Ensure: $\tau^* < \frac{-\log(MeR^{-1})}{C_0}$
Ensure: $\mu < \left(MeR^{-1}\right)e^{C_0\tau^*} < 1$
  $N \leftarrow 1$
  while $N \neq \overline{N}$ **do**
    Find $\overline{A}$ from (16)
    Find $\overline{B}$ from (17)
    $\Theta(N) \leftarrow D\mu^N + \tau^*\overline{A}\overline{z}\overline{B}$
    $N \leftarrow N + 1$
  end while
  if $\min[\Theta(N)] \leq \Delta$ **then**
    $N^* \leftarrow \arg\min_N[\Theta(N)]$
    Return $\hat{A}$ using N^*
  else if $\min[\Theta(N)] > \Delta$ **then**
    Return "Failed"
  end if

VI. Simulation

To demonstrate the effectiveness of the system identification approach in Section III, the Van der Pol Oscillator is used. Consider the Van der Pol Oscillator given by

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 - x_2 + x_2x_1^2. \end{aligned} \tag{19}$$

A total of 209 trajectories of the system in (19) are sampled , each over the time interval $[0, 10]$ with a sampling time of 0.02 seconds. The trajectories are recorded starting form initial conditions that are uniformly sampled from the set $[-1, 1] \times [-1, 1]$. The trajectories are then lifted to different dimensions, with $N = 2, 3, ..., 12$ for system identification.

Using the lifted trajectories, (10) is used to determine the system matrix $\hat{A}$. The identified system matrix is used to simulate the identified system from $t_0 = 0$ to $t_f = 20$ seconds. The trajectories generated by the identified system are truncated to the first two dimensions and compared with the recorded trajectories of the nonlinear system in (19). For comparison, trajectories of the linear system obtained using the model-based Carleman linearization in [9] are also generated. Note that the system matrix $\hat{A}$ is computed directly using recorded data from the nonlinear system. The model-based Carleman linearization from [9] is used for comparison and analysis purposes only.

The resulting trajectories of the nonlinear system, the model-based Carleman-linearized system, and the identified linear systems are shown in Figures 1-3. To quantify the performance, the errors $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\|$ between the measured trajectories and the identified system trajectories, and $\sup_{t\in[0,\tau^*]} \|x(t) - z|_2(t)\|$, between the measured trajectories and the model-based Carleman-linearized trajectories, are plotted in Figure 5. The parameters introduced in Sections II and IV are selected as $M = 1.5$, $R = 4.1$, $C = 33.7$, $\overline{N} = 13$ and $C_0 = 0.001$, all of which satisfy the conditions in Section IV and in (3). Using (18) the selected parameters can be seen to guarantee the error bound $\|x(t) - \hat{z}|_2(t)\| \leq 3$ over the time interval $t \in [0, 0.2]$.

Remark 1. Since the algorithm returns $\hat{A}$ using N^* or returns "Failed", insights into the system identification method and the error bounds for $N \neq N^*$ are not obtainable. To that end, this paper presents results for identified systems using lifting orders other than N^* and compares the guaranteed error bounds and the error for a range of lifting orders.

Figure 4 illustrates the data-driven bound on $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\|$, developed in (18), compared with the model-based bound on $\sup_{t\in[0,\tau^*]} \|x(t) - z|_2(t)\|$, developed in [9].

VII. Discussion

The results in Figures 1-5 show that for the model-based Carleman-linearized system, increasing the order N produces a more accurate approximation of (19). However, as seen in Figure 3 for the data-driven identified system, increasing the order N past a certain point can result in an identified

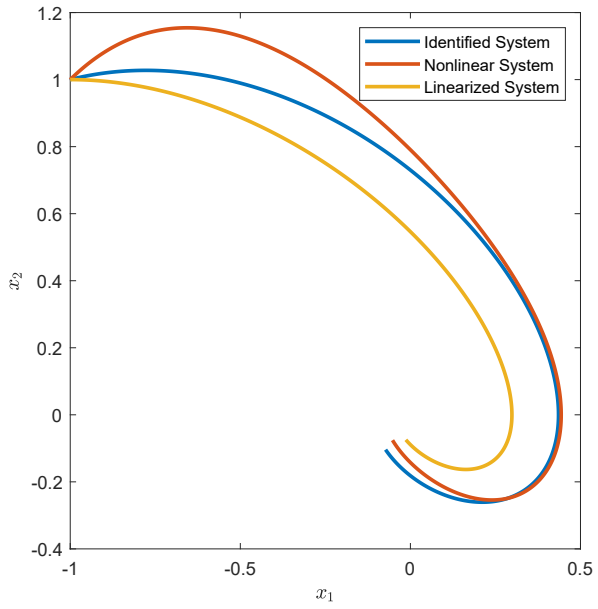

Fig. 1. Trajectories of the nonlinear system, the identified system, and the model-based Carleman-linearized system using $N = 2$.

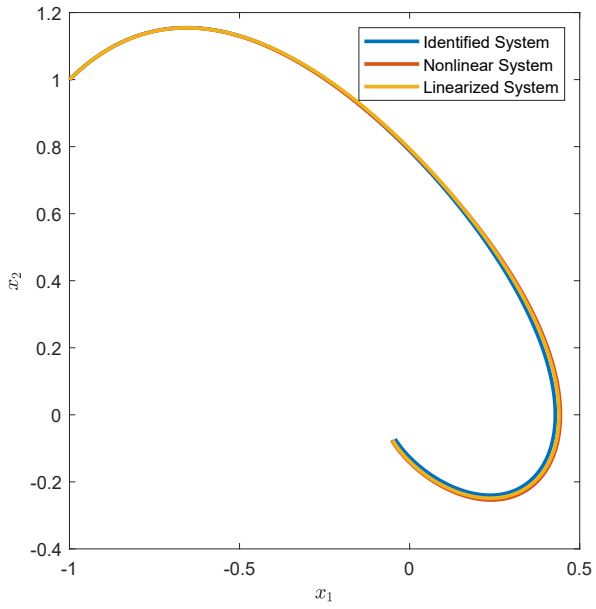

Fig. 2. Trajectories of the nonlinear system, the identified system, and the model-based Carleman-linearized system using $N = 5$.

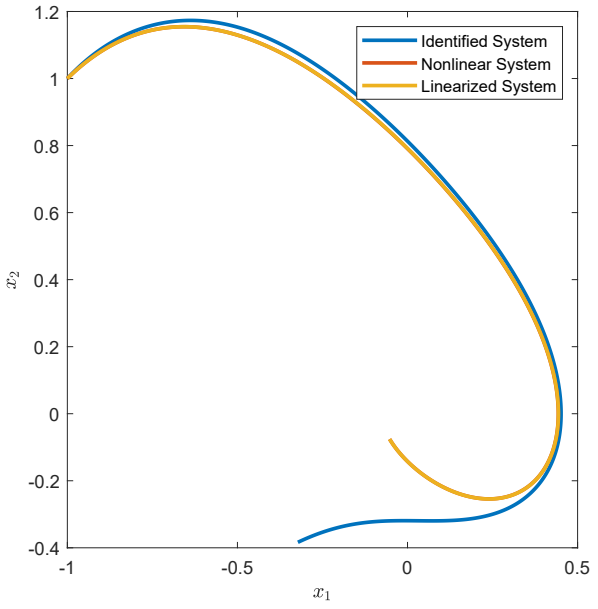

Fig. 3. Trajectories of the nonlinear system, the identified system, and the model-based Carleman-linearized system using $N = 11$.

model that is inaccurate. As indicated by the bound in (18), as long as τ^* is small enough to satisfy (12), the Carleman linearization is guaranteed to get better with increasing truncation order. However, when used in conjunction with a system identification method, there is a critical truncation order after which the estimation becomes less accurate.

The optimal truncation order can be estimated using the bound developed in (18), as illustrated by Figures 5 and 4. Figure 5 shows that in the numerical experiment, $N = 6$ produces the identified system with the lowest error, while the analytical bound, (see Figure 4) indicates that $N^* = 3$, meaning for $N = 3$ the smallest guaranteed error bound is obtained. Since the analytical bound is conservative, some discrepancy between the numerical results and the analytical bound is expected.

To further explore the discrepancy, it is instructive to plot the error between the full trajectories of the model-based Carleman-linearized system and the data-driven identified linear system (see Figure 6) and the error between the same trajectories, truncated to the first two dimensions (see Figure 7). The results indicate that while the full state estimation error $\sup_{0\leq t\leq \tau^*} \|z(t) - \hat{z}(t)\|$ increases monotonically with N, the truncated state estimation error $\sup_{0\leq t\leq \tau^*} \|z|_2(t) - \hat{z}|_2(t)\|$ initially decreases and then increases with increasing N. The analysis presented in this paper, that produces $N^* = 3$, is based on a bound on the full state estimation error, which is observed to be an overly conservative bound on the truncated state estimation error for a subset of truncation orders. The authors postulate that development of an error bound for the truncated state estimation error will reduce the discrepancy between the analytical and the experimental results.

VIII. Conclusion and Future Work

A system identification method using Carleman linearization is developed to identify a lifted linear system that produces trajectories which remain within a guaranteed error bound from the trajectory of a nonlinear system under mild

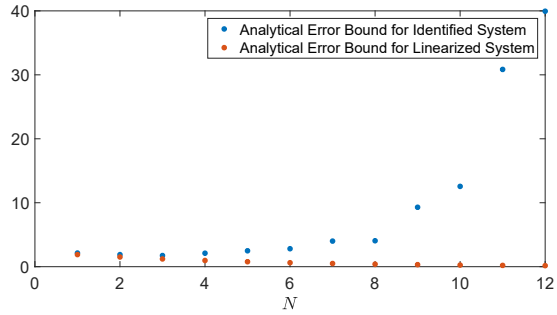

Fig. 4. Comparison of analytical errors bounds on $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\|$, the error between the trajectories of the nonlinear system and the identified system, and $\sup_{t\in[0,\tau^*]} \|x(t) - z|_2(t)\|$, the error between the trajectories of the nonlinear system and the model-based Carleman-linearized system, for truncation orders $N = 1, 2, ..., 12$.

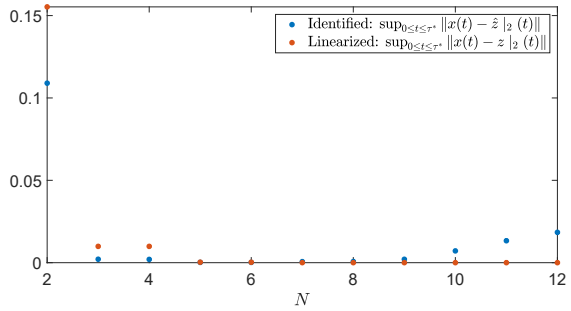

Fig. 5. Comparison of errors $\sup_{t\in[0,\tau^*]} \|x(t) - \hat{z}|_2(t)\|$ between the measured trajectories and the identified trajectories, and $\sup_{t\in[0,\tau^*]} \|x(t) - z|_2(t)\|$, between the measured trajectories and the model-based Carleman-linearized trajectories, for truncation orders $N = 1, 2, ..., 12$.

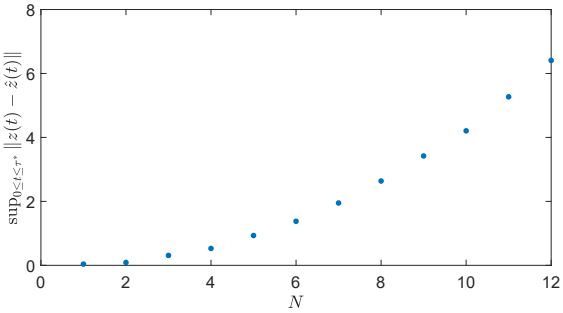

Fig. 6. maximum error norm of the full state between the linearized system and the identified system using $N = 1, 2, ..., 12$

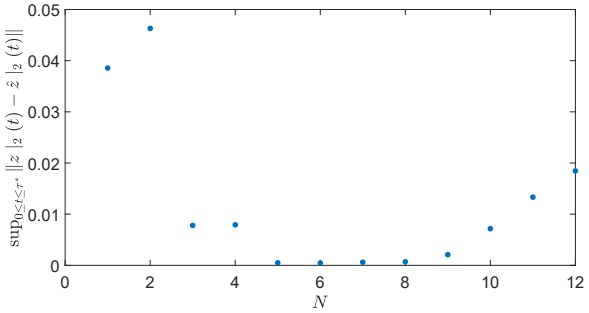

Fig. 7. maximum error norm of the first two states between the linearized system and the identified system using $N = 1, 2, ..., 12$

assumptions. The effectiveness of the technique is demonstrated by identifying an approximation of the Van der Pol oscillator from data.

The time interval $[0, \tau^*]$ over which the error bound is guaranteed is determined by the decay rate in Assumption 2. In practice, it is preferable for τ^* to be dictated by the user. While data-driven construction of a linear system that approximates trajectories of a nonlinear system up to arbitrary precision, over an arbitrarily large time interval, given a finite database of trajectories is infeasible, recent results such as [12] indicate that model-based construction of such a linear system is possible, albeit for a restricted class of nonlinear systems. Examination of such restricted classes from a data-driven construction perspective is a part of future research. Future research will also focus on analytical estimation of the truncated state estimation error, the bound in Assumption 4, and the effect of measurement noise on the error bound and investigation of machine precision effects at high truncation orders.

REFERENCES

[1] L. Ljung, *System identification — theory for the user*. Prentice Hall, 1999.

[2] M. N. Mahyuddin, J. Na, G. Herrmann, X. Ren, and P. Barber, "Adaptive observer-based parameter estimation with application to road gradient and vehicle mass estimation," *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2851–2863, 2013.

[3] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 113, no. 15, pp. 3932–3937, 2016.

[4] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, "A kernel-based method for data-driven Koopman spectral analysis," *J. Comput. Dyn.*, vol. 2, no. 2, pp. 247–265, 2015.

[5] J. A. Rosenfeld, R. Kamalapurkar, B. Russo, and T. T. Johnson, "Occupation kernels and densely defined Liouville operators for system identification," in *Proc. IEEE Conf. Decis. Control*, Dec. 2019, pp. 6455–6460.

[6] T. A. Tutunji, "Parametric system identification using neural networks," *Applied Soft Computing*, vol. 47, pp. 251–261, 2016.

[7] G. Mamakoukas, M. L. Castaño, X. Tan, and T. D. Murphey, "Derivative-based koopman operators for real-time control of robotic systems," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 2173–2192, 2021.

[8] T. Carleman, "Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires," *Acta Mathematica*, vol. 59, pp. 63–87, 1932.

[9] A. Amini, Q. Sun, and N. Motee, "Error bounds for carleman linearization of general nonlinear systems," in *2021 Proceedings of the Conference on Control and its Applications*. SIAM, 2021, pp. 1–8.

[10] E. A. Coddington, *An introduction to ordinary differential equations*. Courier Corporation, 2012.

[11] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," *SIAM Review*, vol. 23, no. 1, pp. 53–60, 1981.

[12] A. Amini, C. Zheng, Q. Sun, and N. Motee, "Carleman linearization of nonlinear systems and its finite-section approximations," *arXiv preprint arXiv:2207.07755*, 2022.